On Processing Relative Clauses in Head-final Languages: Evidence of coping mechanisms for working memory

Kris Ramonda

Abstract

Although most of the world's head-initial languages have a noun + relative clause constituent order, head-final languages are inconsistent in the syntactical positioning of relative clauses. This inconsistency could be due, in part, to competing elements between the heavy constituent principle (Hawkins, 1994) and Lehmann's (1974) basic constituent order predictions. This study examines possible causes of head-final relative clause syntactical inconsistencies, how these inconsistencies impact working memory, and what potential coping mechanisms compensate for the heavy cognitive load of relative clause + noun constituent order. A comparative analysis taken from samples of relative clauses translated by native speakers into nine different languages (five noun + relative clause languages and four relative clause + noun languages) served as the basis for the study. Results indicate that relative clause + noun constituent order seems to place more restrictions on the types of relative clauses permissible in the language.

Key terms: working memory, relative clauses, head-final languages, the heavy constituent principle

Relative clauses, because of the large number of grammatical and lexical elements that they contain, require a significant amount of cognitive resources to process, which in turn places a higher burden on working memory. According to the *heavy constituent principle* (Hawkins, 1994), there is a tendency for languages to place process-heavy constituents after the head noun in relative clauses. As a result, the majority of the world's languages are **noun + RC** constituent order. However, as Lehmann's (1973) basic constituent order correlations predict, OV languages which are head-final, tend to be **RC + noun**. As a result of these competing principles, the RC/noun constituent order of OV languages is inconsistent. If an OV language has an **RC + noun** constituent order, then this would violate the *heavy constituent principle*. Violation of this principle could possibly entail a heavier processing load for speakers of those languages. It is conceivable, however, that languages which violate the *heavy constituent principle* utilize strategies or coping mechanisms in order to compensate for the additional burden on working memory.

Currently, no known study has been carried out to compare and contrast the organization of relative clauses in **RC + noun** and **noun + RC** languages with consideration for both working memory and the *heavy constituent principle*. As such, the purpose of this study is to analyze the syntactical and morphological differences found between these two categories of language to investigate whether any specific strategies are employed by the **RC + noun** languages in order to facilitate the heavier processing burden.

Background

Many aspects of relative clauses have been investigated with respect to their impact on working memory. Since limitations on working memory are related to the complexity of incoming input, it logically follows that second language learners would bear more of an input processing burden when compared to the ease with which they process L1 input. In one study, (Juffs, 2007) it was suggested that processing difficulties of relative clauses for second language learners are linked to the *Noun Phrase Accessibility Hierarchy*, NPAH (Keenan & Comrie, 1977).

Noun Phrase Accessibility Hierarchy

SUBJECTS > DIRECT OBJECTS > INDIRECT OBJECTS > OBLIQUES > GENITIVES > OBJECT OF COMPARISON

The NPAH predicts what types of relative clauses a language will have by specifying an ordered hierarchy in which having a type of relative clause on any position in the hierarchy would implicate that the same language would also have all other positions to the left as well. If it is assumed the positions further to the left represent the more ubiquitous, process-friendly of the relative clauses, then for individuals this could represent certain stages of development of the interlanguage. Since the left-most type of RC is subject relative clause, this might be the first kind of RC acquired in the acquisition orders of relative clauses for second language learners.

Similar conclusions about comparative difficulty of RCs within the NPAH have been supported by a number of studies which have shown that subject relative clauses are more easily processed than object RCs for L1 speakers (Blaubergs & Braine, 1974; Brown,& Hagoort, 2000; Ford, 1983; King & Just, 1991; Schriefers et al., 1995; Traxler et al., 2002; Traxler, et al., 2005). The relative processing ease of subject relative clauses over object relative clauses has been attributed to a number of factors. King and Just (1991) proposed that subject RCs are easier to process because they have a canonical word structure, whereas object relative clauses oftentimes do not (see Appendix II for further explanation). The implication is that changes in the underlying word order of a given language could negatively impact processing speed, even if a particular structure (object relative clauses in this case) necessitates such a change.

Some alternative hypotheses have been introduced as a means of explaining the relative processing difficulty of non-subject relative clauses. The *active filler hypothesis*, AFH, (Clifton & Frazier, 1989) predicts that a relative pronoun is assumed to be a subject until otherwise detected. In the case that it is not a subject, the parser must reanalyze the relative pronoun. If the processing burden of non-subject relative clauses is approached from the perspective of the AFH, then it would follow that object, oblique, possessive, and comparative relative clauses would all entail a heavier processing burden since they would have to be reanalyzed by the parser. The AFH would not, though, predict increasingly greater degrees of processing burden as positions move further and further to the right of the NPAH, since only the subject relative clause would be spared reanalysis.

Despite these findings which support the notion that subject relative clauses are easier to process than other types, there is some evidence of the contrary. One study of relative clauses in Chinese (Chen et al. 2008) found that the processing of object relative clause sentences placed fewer demands on working memory compared to that of subject relative clauses sentences. This was attributed to the observation that processing the subject relative clause sentence requires more cognitive resources from working memory than does processing the object relative clause sentences due to the number of syntactical heads they require.

If it is assumed, however, that the distance of dependencies (this refers to the distance between two related lexical elements) is a function of difficulty of processing RCs, then the case of Chinese is highly unique. As Lin (2008) notes, distance of dependency patterns are opposite between languages when each language demonstrates identical basic word order but opposing head positions (head-initial or head-final). As it happens, Chinese is an anomaly in this sense, because it is an SVO language which has an **RC + noun** constituent order. In fact, Li (1994) claims that Chinese is the only known language which is both SVO and **RC + noun** word order (presumably unique because it violates both the *heavy constituent principle* and Lehmann's 1973 *Constituent Order Correlations*). Given this, the findings from the aforementioned study on comparative ease of object relative clause processing in Chinese can be better understood by considering the opposing distance of dependency patterns. Such a pattern might not contribute an additive burden on processing of object relative clauses due to the particular combination of linguistic features of the Chinese language.

Other than Chinese, there are a number of additional East Asian languages which share some commonalities with the way RCs are structured, such as having fewer cues to signal relativization when compared to Indo-European languages (Comrie,

2007). Japanese is an example of this, which seems not to exhibit any sort of overt marker for relativization. Moreover, some researchers have gone as far as to suggest that the supposed RCs in some of these East Asian languages could be seen as adjuncts and not RCs at all (Lin 2008; Hawkins 2007) because they have no dependency to the head noun. The current study seeks to take a working memory-based perspective to draw on differences such as these and others which may surface in the emergent data among the **noun + RC** and **RC + noun** languages sampled.

Method

The current study has sampled five **noun + RC** and four **RC + noun** languages to investigate what, if any, additional strategies emerge in the **RC + noun** languages since these languages violate the *heavy constituent principle*. The research questions are the following:

1. Do **RC + noun** languages employ any particular strategies or coping mechanisms to ease the processing burden of relative clauses when compared to **noun + RC** languages?
2. If **RC + noun** languages have particular coping strategies, what are they?

As for the sample of languages chosen for this study, the **noun + RC** languages included English, German, Russian, Arabic, and Spanish; the **RC + noun** languages included Japanese, Chinese, Uzbek, and Turkish. The translations which resulted in the data presented in this study were elicited from native speakers, and because of this, the languages selected were limited to those for which native speakers were accessible. In this case, the native speakers were all international graduate students enrolled in an MA program of applied linguistics at a southwestern university in the USA. The native speaker participants were given a series of sentences in English which contained different types of relative clauses and were asked to translate the sentences into their L1. Participants were given as much time as they needed to complete the translations and the researcher assisted in explaining the meaning of a sentence if there were any doubts about the relative clause construction.

There were a total of four sentences consisting of subject, object, oblique, and possessive relative clause constructions. The following were the English sentences that the native speaker participants were asked to translate:

1. John is the guy who likes Mary.
2. John is the guy who Mary likes.
3. John is the guy whom Mary gave money to.
4. John is the guy whose family Mary met.

After gathering all the data, the sentences were aligned, color coded, and grouped according to their noun/RC ordering so that they could be analyzed more easily (see Appendices I-IV). The head nouns and relativizers were color coded for all sentences and were used to visually observe any salient features or omissions of features as well as any interesting syntactical patterns among and across the two groups. Although this method of looking for patterns was largely inductive, a number of questions guided the analysis and served as points of observation, such as:

- What types of RCs are permissible in the sample of **RC + noun** vs. **noun + RC** languages?
- What syntactical operations are involved in each type of RC?
- What type of relativizer, if any, is used, and where is it located?

Results

As for the types of RCs which were permissible in the sample of languages for this study, the data indicate that the **Noun + RC** languages allow for all four types (subject, object, oblique, and possessive). For the **RC + noun** language group, however, Japanese and Chinese did not seem to allow for a possessive RC construction.

(1) CHINESE (RC + N)

John shi na ge ren, ta de jiaren he Mary jianmian.
[John be dem class person he gen family conj Mary meet]

(2) JAPANESE (RC + N)

Mary ga a-tta kazoku wa John no kazoku desu.
[Mary nom meet-pst family top John gen family be]

Figures (1) and (2) are translations of the English sentence, "John is the guy whose family Mary met." As can be seen in figure (1), two independent clauses are required to construct a possessive relative clause. In this particular case, two additional native speakers of Chinese were consulted to confirm this. One of those speakers offered an alternative construction that would allow for the possessive relative clause without having to combine two independent clauses.

(3) CHINESE (RC + N)

John jiu shi na ge jiaren he Mary jianmian de ren.
[John just be dem class family conj Mary meet rel person]

In figure (3) notice how unlike figure (1) a relativizer is used and it is self-contained within a single independent clause. However, it should be noted that all three consultants suggested that although this sentence is grammatically sound, it probably would not be constructed in such a way by a native speaker and sounded "wordy" as if it had been "translated from another language into Chinese". In other words, despite the fact that this sort of construction is possible in Chinese, it seems to be, at least from this limited evidence, a structure which is avoided by native speakers.

In figure (2), again we seem to see a strategy which involves the use of two clauses to express the meaning of the phrase in English, "John is the guy whose family Mary met". The "wa" topic marker appears to be modifying the entire first clause so the literal translation would be something like "As for the family Mary met, it is John's family". In this instance it is apparent that there is not any relative clause at all. This data, therefore, supports the notion that possessive relative clauses in Japanese are non-existent, and that in Chinese, while possible, seem to be avoided by native speakers because of the unnaturalness of the construction.

With regards to the syntactical operations employed across the two groups, there are also some notable differences which appeared in the data. For the **noun + RC** group, the data revealed that when comparing subject and object RCs, the languages encoded this difference either by switching the order of the noun and verb at the end of the sentence, or having morphological distinctions embedded in the relativizer or verb.

(4) ARABIC (N + RC)

John howa el rajol elathe y oheb Mary.
[John he art man rel msc like Mary]
John is that guy who likes Mary.

(5) ARABIC (N + RC)

John howa er rojol elathe Mary to heb.
[John he art man rel Mary fem like]

John is the guy who Mary likes.

(6) GERMAN (N + RC)

John ist der Mann, der Mary mag.
[John be art guy, rel Mary like]
John is the guy who likes Mary.

(7) GERMAN (N + RC)

John ist der mann, den Mary mag.
[John be art guy, rel Mary like]
John is the guy who Mary likes.

The **RC + noun** languages, however, seem to show a preference for case markings instead of syntactical movement (although Chinese appears to be the exception to this) as can be seen in the following examples:

(8) UZBEK (RC + N)

John Mary ni yaxshiko'ruvchgan bola.
[John Mary art like guy.]
John is the guy who likes Mary.

(9) UZBEK (RC + N)

John Mary yaxshiko'radigan bola.
[John Mary like guy]
John is the guy who Mary likes.

(10) CHINESE (RC + N)

John shi xihuan Mary de na ge ren.
[John be like Mary Rel Dem class person]
John is the guy who likes Mary.

(11) CHINESE (RC + N)

John shi Mary xihuan de na ge ren.
[John is Mary like rel dem class person]
John is the guy who Mary likes.

Finally, the use and syntactical position of the relativizer was examined. As for the **noun + RC** group, a relativizer was used for every language in the sample, and its position was always immediately following the head noun, as can be seen in figure (12).

(12) SPANISH (N + RC)

John es el tipo quien gusta de Mary.

[John be art guy Rel like Prep Mary]

As for the **RC + noun** language group, the relativizer was utilized in a number of ways. Chinese was most similar to the **noun + RC** language group in that it had an independent overt marker, the character 的 (de). Unlike the **noun + RC** groups, though, the relativizer was not positioned adjacent to the head noun.

CHINESE (RC + N)

John	shi	Mary	gei	qian	de	na	ge	ren.
[John	be	Mary	give	money	rel	dem	class	person]

Uzbek and Turkish also utilized a relativizer, but it was encoded within the verb.

TURKISH (RC + N)

Mary-nin	para	ver-digi	adam	John.
[Mary-poss	money	give-pstpart(rel)	guy	John]

Japanese does not supply any type of relativizer at all.

JAPANESE (RC + N)

John	ga	Mary	kara	okane	o	mor-atta	otokonohito	desu.
[John	Nom	Mary	prep	money	acc	pass-pst	guy	is]

Discussion

The first point of observation from the data revealed that the **noun + RC** languages all allowed for possessive RCs, but this was not the case for all of the **RC + noun** languages. If we assume, as Juffs (2007) suggested, that the NPAH is linked to processing difficulties, then it would follow that **RC + noun** languages would be less likely to accommodate those RCs further to the right of the NPAH. More specifically, Japanese did not allow for a possessive relative clause construction at all. In Chinese, although it was grammatically possible, it was rejected as being an unnatural response by all three native speaker consultants. It could be that this kind of construction never came about in Chinese because the additional possessive element adds to the processing burden, making it easier just to break it into separate clauses rather than resorting to a possessive relative clause to express it.

The syntactical operations between both groups did not seem to reveal any particular dominant strategies for either group, although the **RC + noun** group, with the exception of Chinese, did not use any inversion of the verb and noun to make the distinction between subject and object relative clauses. Given that Chinese is somewhat of an anomaly in the way it treats RCs, it would be beneficial to increase the sample size of **RC + noun** languages to see if other languages conform to this tendency of noun/verb inversion. Even if this tendency is observable across other **RC + noun** languages, however, it is unclear what possible, if any, coping mechanism this might afford to facilitate comprehension of RCs.

The last point of observation, the existence and positioning of the relativizer, led to some interesting contrasts between the two groups. One salient feature of the sample of **noun + RC** languages was that they all had overt relative pronouns, some of which also have different morphological suffixes. Furthermore, for all of the five **noun + RC** languages, the head noun and relative pronouns were adjacent to one another (the one exception was in certain constructions of Spanish in which obligatory prepositions attached to the verbs were positioned between the head noun and relative pronoun). In the **RC + noun** language

group, on the other hand, the data showed a much more erratic distribution. Japanese appeared not to have any relativizer at all, while the other languages within this group had either a distinct relative pronoun (Chinese), or a verbal inflection or case marking to indicate a relative clause.

Conclusion

A number of salient differences between the two groups of languages compared in this study surfaced. The differences in the data between allowable RC types between these two groups provided the most compelling evidence among the three points of observation. If Juff's (2007) proposal about the NPAH and the relative difficulty among types of relative clauses is assumed to be true, then it can be hypothesized that **RC + noun** languages will be less likely than **noun + RC** languages to have relative clause types further to the right of the NPAH. In the current study, there seems to be some evidence to support this, since all five of the **noun + RC** languages are able to relativize on subject, object, oblique, and possessive, yet, half of those **RC + noun** languages were unable to construct a possessive relative clause (Chinese and Japanese) which is the position furthest to the right of the types compared in this study on the NPAH (data for the comparative relative clause was not gathered in this study). Future studies could include a larger sample of languages and include comparative relative clauses to see if a similar tendency emerges.

In view of the data available, one plausible strategy could be the tendency (when compared to **noun + RC** languages) for **RC + noun** languages to eschew those more process-heavy relative clause types that appear to the right of the NPAH. Such a hypothesis would predict that for each position on the NPAH, **RC + noun** languages would, taken as a whole, exhibit stronger aversion to these types of constructions when compared to **noun + RC** languages. The case of Chinese provided interesting anecdotal evidence as well, since, in addition to the two consultants who provided the initial two possible constructions, another was consulted to confirm the unnaturalness of figure (3). More specifically, the strategy for possessive RCs in Chinese and Japanese appeared to be to simply separate the English sentence, "John is the guy whose family Mary met" into two separate clauses, thereby restructuring it to be more process-friendly. Future studies are necessary in order to confirm or reject the implications of these findings by sampling and contrasting a larger number of languages in each respective group.

Limitations

The sample size of languages used was much too small to draw any strong conclusions and some of the languages used in the study belonged to the same language family. Due to this, the findings can serve at best as a starting point for future studies which can compare and contrast a greater pool of languages and constructions from a variety of language phyla. Furthermore, comparative relative clauses were not included in this particular study, but their inclusion would make for a more complete comparison, particularly since comparative relative clauses are in the right-most position of the NPAH.

References

Blaubergs, M., & Braine, M. (1974). Short-term memory limitations on decoding self-embedded sentences. *Journal of Experimental Psychology*, 102, 745-748.

Brown, C. M., & Hagoort, P. (2000). On the electrophysiology of language comprehension:Implications for the human language system. In M. W. Crocker & M. J. Pickering (Eds.), *Architectures and mechanisms for language processing* (pp. 213–237). NewYork: Cambridge University Press.

Chen, B., Ning, A., Bi, H., & Dunlap, S. (2008). Chinese subject-relative clauses are more difficult to process than object-relative clauses. *Acta Psychologica*, 129, 61-65.

Comrie, B. (2007). The acquisition of relative clauses in relation to language typology. *Studies in Second Language Acquisition*, 29, 301-309.

Clifton, C., & Frazier, L. (1989). Comprehending sentences with long-distance dependencies.In G. N. Carlson & M. K.

Tanenhaus (Eds.), *Linguistic Structure in Language Processing* (pp. 273-317). Dordrecht. Kluwer.

Ford, M. (1983). A method for obtaining measures of local parsing complexity throughout sentences. *Journal of Verbal Learning & Verbal Behavior*, 22, 203-218.

Hawkins, J. (2007). Acquisition of relative clauses in relation to language universals. *Studies in Second Language Acquisition*, 29, 337-344.

Hawkins, J. (1994). A Performance Theory of Order and Constituency. Cambridge University Press, Cambridge.

Juffs, A. (2007). Second language acquisition of relative clauses in the languages of East Asia. *Studies in Second Language Acquisition*, 29, 361-365.

Keenan, E., & Comrie, B. (1977). Noun phrase accessibility and universal grammar. *Linguistic Inquiry,* 8, 63-99.

King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory & Language*, 30, 580-602.

Lehmann, W. (1973). A structural principle of language and its implications. *Language*, 49, 47-66.

Li, C., & Li, L. (1994). Relativization in English and Chinese. *Papers and Studies in C Contrastive Linguistics*, 29, 61-75.

Lin, C. (2008). The processing foundations of head-final relative clauses. *Language and Linguistics*, 9, 813-838.

Schriefers, H., Friederici, A.D., & Kuhn, K. (1995). The processing of locally ambiguous relative clauses in German. *Journal of Memory & Language*, 34, 499–520.

Traxler, M., Morris, R., & Seely, R. (2002). Processing subject and object relative clauses: Evidence from eye movements. *Journal of Memory & Language*, 47, 69-90.

Traxler, M., Williams, R., Blozis, S., & Morris, R. (2005). Working memory, animacy, and verb class in the processing of relative clauses. *Journal of Memory & Language*, 53, 204-224.

Appendix I

ENGLISH (N + RC)

John	is	the	guy	who	likes	Mary.
[John	be	ART	guy	Rel	like	OBJ]

SPANISH (N + RC)

John	es	el	tipo	quien	gusta	de	Mary.
[John	be	art	guy	Rel	like	Prep	Mary]

RUSSIAN (N + RC)

John	paren;	kotor-omu	nravit-sa	Mary
[John	guy	rel-dat	like-ref	Mary]

ARABIC (N + RC)

John	howa	el	rajol	elathe	y	oheb	Mary.
[John	he	art	man	rel	msc	like	Mary]

GERMAN (N + RC)

John	ist	der	Mann,	der	Mary	mag.
[John	be	art	guy,	Rel	Mary	like]

JAPANESE (RC + N)

John ga Mary o suki na otokonohito desu.
[John Nom Mary Acc like ptl guy is]

UZBEK (RC + N)

John Mary ni Yaxshiko'ruvchgan bola.
[John Mary art like guy.]

CHINESE (RC + N)

John shi xihuan Mary de na ge ren.
[John be like Mary Rel Dem class person]

TURKISH (RC + N)

Mary-yi begenen adam John.
[Mary-acc like guy John]

Appendix II

ENGLISH (N + RC)

John is the guy who Mary likes.
[John be art guy rel Mary like]

SPANISH (N + RC)

John es el tipo que le gusta a Mary.
[John be art guy rel IOpro like prep Mary]

RUSSIAN (N + RC)

John paren; kotoryi nravit-sa Mary
[John guy; rel like-ref Mary]

ARABIC (N + RC)

John howa er rojol elathe Mary to heb.
[John he art man rel Mary fem like]

GERMAN (N + RC)

John ist der mann, den Mary mag.
[John be art guy, rel Mary like]

JAPANESE (RC + N)

John wa Mary ga suki na otokonohito desu.
[John top Mary Nom like ptl guy is]

UZBEK (RC + N)

John Mary yaxshiko'radigan bola.
[John Mary like guy]

CHINESE (RC + N)

John shi Mary xihuan de na ge ren.
[John is Mary like rel dem class person]

TURKISH (RC + N)

Mary-nin begen-digi adam John.
[Mary-poss like-pstPart guy John]

Appendix III

ENGLISH (N + RC)

John is the guy whom Mary gave money to.
[John be art guy rel Mary give money prep]

SPANISH (N + RC)

John es el tipo a quien Mary le d-io dinero.
[John be art guy prep rel Mary IOpro give-pst money]

RUSSIAN (N + RC)

John paren; kotor-omu Mary da-la dengi
[John guy; rel-dat Mary give-pst money]

ARABIC (N + RC)

John howa er rajol elathe Mary atathou felou.
[John he art man rel Mary give money]

GERMAN (N + RC)

John ist der mann, dem Mary Geld gab.
[John be the guy, rel Mary money gave]

JAPANESE (RC + N)

John ga Mary kara okane o mor-atta otokonohito desu.
[John Nom Mary prep money acc pass-pst guy is]

UZBEK (RC + N)

John Mary pul bergan bola.
[John Mary money gave guy]

CHINESE (RC + N)

John shi Mary gei qian de na ge ren.
[John be Mary give money rel dem class person]

TURKISH (RC + N)

Mary-nin para ver-digi adam John.
[Mary-poss money give-pstpart guy John]

Appendix IV

ENGLISH (N + RC)

John is the guy whose family Mary met.
[John be art guy rel family Mary meet]

SPANISH (N + RC)

John es el tipo cuya familia Mary conoc-ió.
[John be art guy rel family Mary meet-pst]

RUSSIAN (N + RC)

John paren; ch'u sem'u vstret-ila Mary.
[John guy; rel-acc family-acc meet-pst Mary]

ARABIC (N + RC)

John howa er rajol elathe Mary kabal ousrotahon.
[John he art man rel Mary meet family]

GERMAN (N + RC)

John ist der Mann, dessen familie Mary getroffen hat.
[John be art guy, rel family Mary meet aux]

JAPANESE (RC + N)

Mary ga a-tta kazoku wa John no kazoku desu.
[Mary nom meet-pst family top John gen family be]

UZBEK (RC + N)

John Mary tomonidan oilasi ko'rilgan bola.
[John Mary prep family meet guy]

CHINESE (RC + N)

John shi na ge ren, ta de jiaren he Mary jianmian.
[John be dem class person he gen family conj Mary meet]

TURKISH (RC + N)

Mary-nin,	aile-si-yle	tanis-tigi	adam	John.
[Mary-poss	family-poss-ins	meet-pstpart	guy	John]